

A New Approach to Automated Epileptic Diagnosis Using EEG and Probabilistic Neural Network

Forrest Sheng Bao^{1,2}, Donald Yu-Chun Lie², and Yuanlin Zhang¹

¹ Dept. of Computer Science ² Dept. of Electrical and Computer Engineering
Texas Tech University, Lubbock, Texas, USA

Abstract—Epilepsy is one of the most common neurological disorders that greatly impair patient's daily lives. Traditional epileptic diagnosis relies on tedious visual screening by neurologists from lengthy EEG recording that requires the presence of seizure (ictal) activities. Nowadays, there are many systems helping the neurologists to quickly find interesting segments of the lengthy signal by automatic seizure detection. However, we notice that it is very difficult, if not impossible, to obtain long-term EEG data with seizure activities for epilepsy patients in areas lack of medical resources and trained neurologists. Therefore, we propose to study automated epileptic diagnosis using interictal EEG data that is much easier to collect than ictal data. The authors are not aware of any report on automated EEG diagnostic system that can accurately distinguish patients' interictal EEG from the EEG of normal people. The research presented in this paper, therefore, aims to develop an automated diagnostic system that can use interictal EEG data to diagnose whether the person is epileptic. Such a system should also detect seizure activities for further investigation by doctors and potential patient monitoring. To develop such a system, we extract four classes of features from the EEG data and build a Probabilistic Neural Network (PNN) fed with these features. Leave-one-out cross-validation (LOO-CV) on a widely used epileptic-normal data set reflects an impressive 99.5% accuracy of our system on distinguishing normal people's EEG from patient's interictal EEG. We also find our system can be used in patient monitoring (seizure detection) and seizure focus localization, with 96.7% and 77.5% accuracy respectively on the data set.

Index Terms—Epilepsy, Electroencephalogram (EEG), Probabilistic Neural Network (PNN), seizure.

I. INTRODUCTION

EPILEPSY is a chronic disorder characterized by recurrent seizures, which may vary from muscle jerks to severe convulsions [1]. Estimated 1% of world population suffers from epilepsy [2], while 85% of them live in the developing countries [3]. Electroencephalogram (EEG) is routinely used clinically to diagnose, monitor and localize epileptogenic zone [4]. Long-term EEG monitoring can provide 90% positive diagnostic information [5] and thus become a golden standard in epilepsy diagnosis.

Traditional methods rely on experts to visually inspect the entire lengthy EEG recordings of up to 1 week, which is tedious and time-consuming. Therefore, many automated system assisting the diagnosis of epilepsy have emerged [6]–[12]. They could detect abnormal EEG segments related to

seizures so that doctors can quickly view events of interest without having to page through the entire recording [13]. But this approach requires the presence of seizure activities in the EEG data. This tough requirement often leads to very long, even up to 1 week, continuous EEG recording to capture seizure activities because of the difficulty to tell if and when a seizure will occur. The long-term EEG recording can greatly disturb patients' daily lives. Another clinical concern is that very unfortunately, 50-75% epilepsy patients in the world reside in areas lack of medical resources and trained professionals [3], which makes the long-term EEG recording virtually impractical to those people. Therefore, an automated EEG epilepsy diagnostic system would be very valuable if it does not require data from active seizure activities (i.e., ictal) to perform the diagnosis. However, to the authors' best knowledge, we are not aware of any report on automated epilepsy diagnostic system using only interictal EEG data.

In this paper, we aim to develop an automated system that can diagnose epilepsy not only by using ictal EEG data but also interictal data. The diagnosis function of this system will be valuable for patients in areas lack of medical resources, and particularly well-trained personnels. Its capability of seizure detection will be the base of effective monitoring in personal health care and help doctors to do further diagnosis if necessary. In addition to diagnosis and seizure detection, we would also like the system to provide basic information on focus localization which is also an important aspect in diagnosis.

Our system is a Probabilistic Neural Network (PNN) [14] based classifier. Previous research suggests PNN is more suitable for medical applications, since it uses Bayesian strategies, a process familiar to medical decision makers [15]. We adopt PNN for its fast speed, high accuracy and real-time property in updating network structure, as we will explain in Sec. IV. It is very difficult to directly use raw EEG data as input to an Artificial Neural Network [16]. Therefore, a key in designing the PNN classifier is to find proper features from the given EEG data and feed those feature values to the classifier (i.e., parameterized EEG data as input). Artificial Neural Network has been used for automated diagnosis by several research groups [9]–[12]. But those work focus on seizure detection only. Since the interictal EEG does not have seizure activities, the features identified for those neural networks might not work for our purpose: diagnosing epilepsy and localizing foci from interictal data. We use four classes of features, namely, power spectral feature, fractal dimensions, Hjorth parameters, and amplitude statistics.

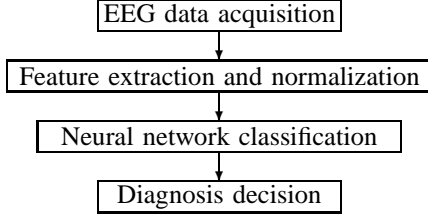


Fig. 1. Flow diagram of our EEG classification scheme

Based on a widely used data set with epileptic and normal EEG data, our experiments indicate that interictal EEG of epileptic patients can be differentiated from those of healthy people with high accuracy and fast speed. Our interictal EEG based diagnostic approach achieves a 99.6% overall accuracy in cross-validation. The existing ictal data based strategy is also tested in our classifier, with 98.3% accuracy. Focus localization is achieved with a 78.5% accuracy. Our classifier is also capable of distinguishing interictal and ictal EEG and thus detecting seizures. The 96.9% accuracy underlines the possible patient monitoring. The speed of our classifier is very fast – 0.01 second per run in all four classification problems. These results imply the possibility of our system for real-time diagnosis, monitoring and focus localization.

II. DATA ACQUISITION

In our experiments, we adopt the data set, which is widely used in previous epileptic diagnosis/analysis research [9], [12], [17], [18], from Klinik für Epileptologie, Universität Bonn, German [17]. It consists of five sets, each containing 100 single-channel EEG segments. Each segment has 4096 sampling points over 23.6 seconds. Note that artifacts, e.g., due to hand or eye movements, have been manually removed by the creators of the data. Data in sets A and B is extracranial EEG from healthy volunteers with eyes open and eyes closed respectively. Sets C and D are intracranial data over interictal period while Set E over ictal period. Segments in D are from within the epileptogenic zone, and those in C from the hippocampal formation of the opposite hemisphere of the brain. All EEG signals were sampled at a sampling rate of 173.61Hz. Refer to [17] for detailed information of the data. The data was filtered by a low-pass filter of cutoff frequency 40Hz.

III. FEATURE EXTRACTION

Our classifier uses 38 features of 4 classes to characterize interictal EEG signal. The power spectral features describe energy distribution in the frequency domain. Fractal dimensions outline the fractal property. Hjorth parameters describe the chaotic behavior. Mean and standard deviation represent the amplitude statistics. Since normalization is very important to distance-based classifier, features are normalized before fed into PNN.

A. Power Spectral Features

To a time series x_1, x_2, \dots, x_N , its Fast Fourier Transform (FFT) X_1, X_2, \dots, X_N are estimated as

$$X_k = \sum_{n=1}^N x_n W_N^{kn}, \quad k = 1, 2, \dots, N$$

where $W_N^{kn} = e^{-\frac{j2\pi kn}{N}}$ and N is the series length.

We noted a clear difference that in general the ictal EEG has more power components in the higher frequency region ($>14\text{Hz}$) while non-ictal EEG are mostly below 14Hz. This point has also been described in [19]. From Fig. 2, it can be clearly seen that the central frequencies of peaks of different EEG signals lie in different regions.

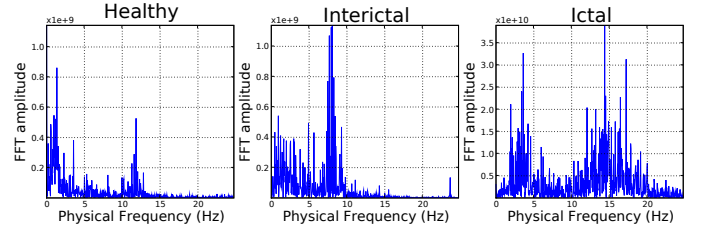


Fig. 2. Typical FFT results of 3 EEG segments (Raw data in μV . For FFT results, the scale in the Y-axis of ictal data is 10 times larger than the ones of healthy and interictal data)

Based on the FFT result, Power Spectral Intensity (PSI) and Relative Intensity Ratio (RIR) are evaluated to each 2-Hz frequency band from 2-32 Hz. The PSI is defined as

$$PSI_k = \sum_{i=\lfloor N \frac{f_{min}}{f_s} \rfloor}^{\lfloor N \frac{f_{max}}{f_s} \rfloor} X_i, \quad k = 1, 2, \dots, 15$$

where $f_{min} = 2k$, $f_{max} = 2k + 2$, f_s is the sampling frequency and N is the series length. As you can see, the f_{min} and f_{max} are the lower and upper boundaries of each 2-Hz band, respectively. The RIR is defined as

$$RIR_j = \frac{PSI_j}{\sum_{k=1}^{15} PSI_k}, \quad j = 1, 2, \dots, 15$$

So we have 15 PSIs and 15 RIPs. They are the first 30 features we used.

The 2-32Hz band covers some EEG abnormalities unique to epilepsy [20], such as the 3Hz and 6Hz spike waves [21].

B. Petrosian Fractal Dimension (PFD)

PFD is defined as:

$$\text{PFD} = \frac{\log_{10} N}{\log_{10} N + \log_{10} \left(\frac{N}{n+0.4N_\delta} \right)}$$

where N is the series length and N_δ is the number of sign changes in the signal derivative [22]. According to Fig. 3 (a), PFD is highly concentrated within each class and there is no overlap among the data for each class either. Therefore, all classes can be clearly distinguished using PFD.

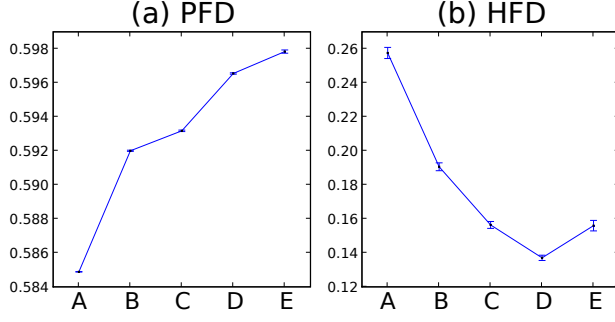


Fig. 3. Average Petrosian Fractal Dimension (a) and Higuchi Fractal Dimension (b) for each set. Error bars denote one standard deviation (σ) for each class of EEG.

C. Higuchi Fractal Dimension (HFD)

Higuchi's algorithm [23] constructs k new series from the original series x_1, x_2, \dots, x_N by

$$x_m, x_{m+k}, x_{m+2k}, \dots, x_{m+\lfloor \frac{N-m}{k} \rfloor k} \quad (1)$$

where $m = 1, 2, \dots, k$.

For each time series constructed from (1), the length $L(m, k)$ is computed by

$$L(m, k) = \frac{\sum_{i=2}^{\lfloor \frac{N-m}{k} \rfloor} |x_{m+ik} - x_{m+(i-1)k}| (N-1)}{\lfloor \frac{N-m}{k} \rfloor k}$$

The average length $L(k)$ is computed as

$$L(k) = \frac{\sum_{i=1}^k L(i, k)}{k}$$

This procedure repeats k_{max} times for each k from 1 to k_{max} , and then uses a least-square method to determine the slope of the line that best fits the curve of $\ln(L(k))$ versus $\ln(1/k)$. The slope is the Higuchi Fractal Dimension. In this paper, k_{max} is 5.

Fig. 3 (b) indicates HFD is also densely clustered within each class and there is a small overlap between classes C and E. HFD, therefore, is a good feature to characterize classes A, B and E.

D. Hjorth Parameters

To a time series x_1, x_2, \dots, x_N , the Hjorth mobility and complexity [24] are respectively defined as

$$\sqrt{\frac{M2}{TP}}$$

and

$$\sqrt{\frac{M4 \cdot TP}{M2 \cdot M2}}$$

where $TP = \sum x_i/N$, $M2 = \sum d_i/N$, $M4 = \sum (d_i - d_{i-1})^2/N$ and $d_i = x_i - x_{i-1}$.

According to Fig. 4 (a), Hjorth mobility has a tight distributions within each class. Even though the Hjorth complexity appears very inconsistent among classes, since PNN uses normalized features, we compute the normalized Hjorth

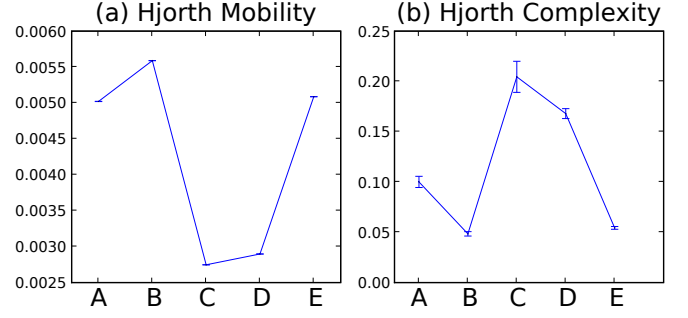


Fig. 4. Average Hjorth Mobility (a) and normalized Hjorth Complexity (b) for each set. Error bars denote one standard deviation (σ) for each class of EEG.

complexity and find it also has a tight distribution within each class.

An interesting finding is that a set with low Hjorth mobility would have high normalized Hjorth complexity. For example, Sets C and D have the lowest mobility and highest complexity, which separates them from other sets.

E. Means and standard deviations

EEG signals from different conditions have different amplitudes. For example, the amplitude of normal activities ranges around $100\mu V$ in this paper while the ictal EEG ranges around $1000\mu V$. Means and standard deviations of both original data and the absolute values of EEG are evaluated. They are the last 4 features we used.

IV. PROBABILISTIC NEURAL NETWORK

An AI-based classifier is essentially a mapping $f: \mathbb{R}^m \rightarrow \mathbb{Z}^n$ from the feature space to the discrete class space. An Artificial Neural Network (ANN) implements such a mapping by using a group of interconnected artificial neurons simulating human brain. An ANN can be trained to achieve expected classification results against the input and output information stream, so there is not a need to provide a specified classification algorithm.

PNN is one kind of distance-based ANNs, using a bell-shape activation function. This technique makes decision boundaries nonlinear and hence it can approach the Bayesian optimal [25]. Compared with traditional back-propagation (BP) neural network, PNN is considered more suitable to medical application since it uses Bayesian strategy, a process familiar to medical decision makers [15]. The real-time property of PNN is also crucial to our research. In PNN, decision boundaries can be modified in real-time as new data become available [14]. There is no need to train the network over the entire data set again. So we can quickly update our network as more and more patients' data becomes available.

Our PNN has three layers: the Input Layer, the Radial Basis Layer which evaluates distances between input vector and rows in weight matrix, and the Competitive Layer which determines the class with maximum probability to be correct. The network structure is illustrated in Fig. 5, using symbols and notations in [26]. Dimensions of arrays are marked under their names.

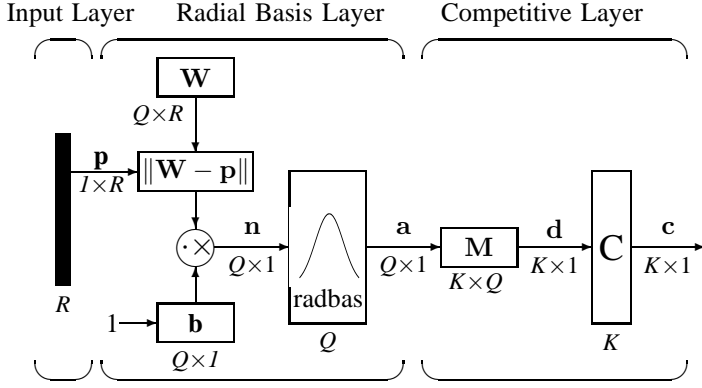


Fig. 5. PNN structure. R : number of features, Q : number of training samples, K : number of classes. The input vector \mathbf{p} is presented here as a black vertical bar.

A. Radial Basis Layer

In Radial Basis Layer, the vector distances between input vector \mathbf{p} and the weight vector made of each row of weight matrix \mathbf{W} are calculated. Here, the vector distance is defined as the dot product between two vectors [25]. The dot product between \mathbf{p} and the i -th row of \mathbf{W} produces the i -th element of the distance vector matrix, denoted as $\|\mathbf{W} - \mathbf{p}\|$. The bias vector \mathbf{b} is then combined with $\|\mathbf{W} - \mathbf{p}\|$ by an element-by-element multiplication, represented as “ \times ” in Fig. 5. The result is denoted as $\mathbf{n} = \|\mathbf{W} - \mathbf{p}\| \times \mathbf{b}$.

The transfer function in PNN has built into a distance criterion with respect to a center. In this paper, we define it as

$$\text{radbas}(n) = e^{-n^2} \quad (2)$$

Each element of \mathbf{n} is substituted into (2) and produces corresponding element of \mathbf{a} , the output vector of Radial Basis Layer. We can represent the i -th element of \mathbf{a} as

$$\mathbf{a}_i = \text{radbas}(\|\mathbf{W}_i - \mathbf{p}\| \times \mathbf{b}_i) \quad (3)$$

where \mathbf{W}_i is the i -th row of \mathbf{W} and \mathbf{b}_i is the i -th element of bias vector \mathbf{b} .

1) *Radial Basis Layer Weights*: Each row of \mathbf{W} is the feature vector of one training sample. The number of rows equals to the number of training samples.

2) *Radial Basis Layer Biases*: All biases in radial basis layer are set to $\sqrt{\ln 0.5}/s$ resulting in radial basis functions that cross 0.5 at weighted inputs of $\pm s$, where s is the spread constant of PNN. In this paper, s is set to 0.1, since our experiments show the highest accuracy is achieved when $s = 0.1$, as illustrated in Fig. 6.

B. Competitive Layer

There is no bias in Competitive Layer. In this layer, the vector \mathbf{a} is first multiplied by layer weight matrix \mathbf{M} , producing an output vector \mathbf{d} . The competitive function \mathbf{C} produces a 1 corresponding to the largest element of \mathbf{d} , and 0's elsewhere. The index of the 1 is the class of the EEG segment. \mathbf{M} is set to $K \times Q$ matrix of Q target class vectors. If the i -th sample in training set is of class j , then we have a 1 on the j -th row of i -th column of \mathbf{M} .

TABLE I
OVERALL ACCURACY AND CLASSIFICATION TIME USING PNN

No.	Experiment	Accuracy	Time (s)
1	normal (200 samples) vs. interictal (200 samples)	99.5%	0.01
2	normal (200 samples) vs. ictal (100 samples)	98.3%	0.01
3	interictal (200 samples) vs. ictal (100 samples)	96.7%	0.01
4	epileptogenic zone (100 samples) vs. opposite hemisphere (100 samples)	77.5%	0.01

V. EXPERIMENTAL RESULTS

As shown in Table. I, we designed four experiments to test the ability of our classifier to separate:

- 1) normal EEG (sets A and B) and interictal EEG (sets C and D)
- 2) normal EEG (sets A and B) and ictal EEG (set E)
- 3) interictal EEG (sets C and D) and ictal EEG (set E)
- 4) interictal EEG sampled from epileptogenic zone (set C) and interictal EEG sampled from opposite hemisphere (set D)

The first two experiments evaluate the performance of our algorithm using interictal EEG and ictal EEG respectively.

The last two experiments evaluate the feasibility of our algorithm on seizure monitoring and focus localization, respectively.

The classifier is validated using leave-one-out cross-validation (LOO-CV) on 400, 300, 300 and 200 samples respectively in experiments 1, 2, 3 and 4. Our algorithm is implemented using the MATLAB Neural Network Toolbox. Table I lists the overall accuracy and classification time of four experiments.

The spread constant of PNN, is selected according to overall accuracy. As illustrated in Fig. 6, all experiments achieve the highest accuracy, when spread constant is 0.1. In our experiments, therefore, spread constant is set to 0.1.

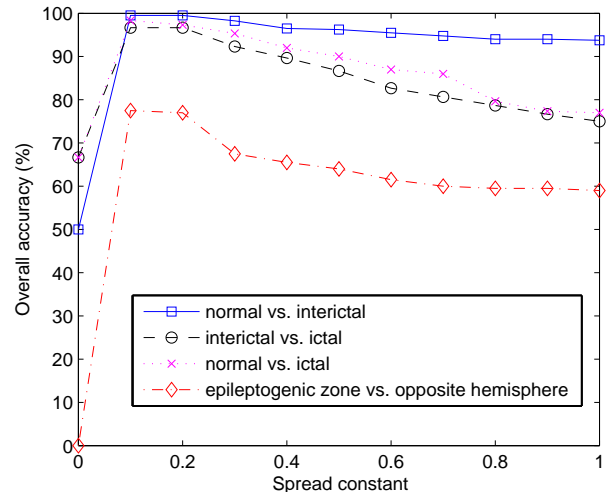


Fig. 6. Overall accuracy with respect to PNN spread constant

Our approach reaches 99.5% accuracy in using interictal EEG for epileptic diagnosis (Experiment 1). This result validates the feasibility to use only interictal data in epileptic diagnosis.

The accuracy of our classifier using ictal data reaches 98.3%, which is very similar to what has been reported from previous work [12].

In Experiment 3, 96.7% accuracy shows our system can distinguish ictal EEG from interictal EEG very well. This suggests the feasibility to continuously monitor patient status or detect seizures by classifying an EEG segment cut out from monitoring data. If the segment covering current time instant is classified as ictal, then the patient has been in an ictal state, i.e., the seizure is occurring. The focus localization experiment achieves a promising accuracy of 77.5%, which still needs further improvement.

In all the 4 experiments, the classification time per run is 0.01 second (on MATLAB R2008a for Linux, 1.6GHz 64-bit CPU, 2G RAM), which is very short compared with the EEG segment length, 23.6 seconds. This shows the feasibility of real-time monitoring. For long-term monitoring, we can periodically sample the EEG by a sliding window and analyze the windowed segment. For example, immediately after ictal activity detection, devices equipped with our algorithm can send out an alarm to healthcare providers.

VI. CONCLUSIONS

In this paper, an automated EEG recognition system for epilepsy diagnosis is developed and validated by cross-validation. Compared with the existing conventional seizure detection algorithms, our approach does not require seizure activity to be captured in EEG recording and thus is seizure-independent. This feature relieves the difficulties in EEG collection since interictal data is much easier to be collected than ictal data. 38 EEG features are extracted and PNN is employed to classify those features.

Experiments indicate that interictal EEG of epileptic patients can be differentiated from those of healthy people with high accuracy and fast speed. Our interictal EEG based diagnostic approach achieves a 99.5% overall accuracy in cross validation. Diagnosis based on ictal data is also tested in our classifier, reaching a high 98.3% accuracy. So, our algorithm works fine with both interictal and ictal data. We also extend the function of the classifier, to patient monitoring and focus localization. 96.7% accuracy is achieved on differentiating ictal from interictal EEG, which suggests the feasibility of online patient monitoring. The focus localization result is also promising with a 77.5% accuracy. The speed of our classifier is very good, costing only 0.01 second to classify an EEG segment of 23.6 seconds.

REFERENCES

- [1] H. Gastaut, *Dictionary of Epilepsy. Part I: Definitions*. World Health Organization, 1973.
- [2] K. Lehnertz, F. Mormann, T. Kreuz, R. Andrzejak, C. Rieke, P. David, and C. Elger, "Seizure prediction by nonlinear eeg analysis," *IEEE Engineering in Medicine and Biology Magazine*, 2003.
- [3] *Atlas: Epilepsy Care in the World*. World Health Organization, 2005.
- [4] R. Flink, B. Pedersen, A. B. Guekht, K. Malmgren, R. Michelucci, B. Neville, F. Pinto, U. Stephani, and C. Ozkara, "Guidelines for the use of EEG methodology in the diagnosis of epilepsy international league against epilepsy," *Acta Neurologica Scandinavica*, vol. 106, 2002.
- [5] C. Logar, B. Walzl, and H. Lechner, "Role of long-term EEG monitoring in diagnosis and treatment of epilepsy," *European Neurology*, vol. 34, 1994.
- [6] A. Gardner, A. Krieger, G. Vachtsevanos, and B. Litt, "One-class novelty detection for seizure analysis from intracranial EEG," *Journal of Machine Learning Research*, vol. 7, 2006.
- [7] V. Srinivasan, C. Eswaran, and N. Sriraam, "Artificial neural network based epileptic detection using time-domain and frequency-domain features," *J. Med. Syst.*, vol. 29, no. 6, pp. 647–660, 2005.
- [8] W. Weng and K. Khorasani, "An adaptive structure neural networks with application to eeg automatic seizure detection," *Neural Network*, vol. 9, no. 7, pp. 1223–1240, 1996.
- [9] V. P. Nigam and D. Graupe, "A neural-network-based detection of epilepsy," *Neurological Research*, vol. 26, pp. 55–60(6), 1 January 2004.
- [10] M. K. Kiymik, A. Subasi, and H. R. Ozcalik, "Neural networks with periodogram and autoregressive spectral analysis methods in detection of epileptic seizure," *J. Med. Syst.*, vol. 28, no. 6, pp. 511–522, 2004.
- [11] N. Pradhan, P. K. Sadasivana, and G. R. Arunodaya, "Detection of seizure activity in EEG by an artificial neural network: A preliminary study," *Computers and Biomedical Research*, vol. 29, no. 4, pp. 303–313, 2002.
- [12] V. Srinivasan, C. Eswaran, and N. Sriraam, "Approximate entropy-based epileptic EEG detection using artificial neural networks," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 3, pp. 288–295, May 2007.
- [13] N. McGorgan. (1999) Neural network detection of epileptic seizures in the electroencephalogram. [Online]. Available: <http://citeseer.ist.psu.edu/625725.html>
- [14] D. Specht, "Probabilistic neural networks," *Neural Networks*, vol. 3, 1990.
- [15] R. K. Orr, "Use of a Probabilistic Neural Network to Estimate the Risk of Mortality after Cardiac Surgery," *Med. Decis. Making*, vol. 17, no. 2, pp. 178–185, 1997.
- [16] C.-W. Ko and H.-W. Chung, "Automatic spike detection via an artificial neural network using raw eeg data: effects of data preparation and implications in the limitations of online recognition," *Clinical Neurophysiology*, vol. 111, no. 3, pp. 477–481, 2000.
- [17] R. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E*, vol. 64, no. 6, p. 061907, Nov 2001.
- [18] I. Güler and E. D. Ubeyli, "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients," *J Neurosci Methods*, vol. 148, no. 2, pp. 113–121, 2005.
- [19] W. Williams, H. Zaveri, and J. Sackellares, "Time-frequency analysis of electrophysiology signals in epilepsy," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 2, pp. 133–143, Mar/Apr 1995.
- [20] M. Sarkela, M. Ermes, M. van Gils, A. Yli-Hankala, V. Jantti, and A. Vakkuri, "Quantification of epileptiform electroencephalographic activity during sevoflurane mask induction," *Anesthesiology*, vol. 107, no. 6, pp. 928–938, 2007.
- [21] J. Alving. EEG and epilepsy. Danish Epilepsy Center. [Online]. Available: [http://www.neuroscienceschool.ku.dk/files/CNM_literature/EEG%20and%20Epilepsy%](http://www.neuroscienceschool.ku.dk/files/CNM_literature/EEG%20and%20Epilepsy%20)
- [22] A. Petrosian, "Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns," in *Proc. of the Eighth IEEE Symposium on Computer-Based Medical Systems*, 1995.
- [23] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Physica D*, vol. 31, no. 2, pp. 277 – 283, 1988.
- [24] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalography and Clinical Neurophysiology*, vol. 29, pp. 306–310, 1970.
- [25] D. Specht, "Probabilistic neural networks for classification mapping, or associative memory," in *Proc. of 1988 IEEE Int'l Conference on Neural Networks*, vol. 1, 1988.
- [26] M. Hagan, H. Demut, and M. Beale, *Neural Network Design*. Thomson Learning, 2002.